UNIVERSITÉ DE MONTRÉAL

SOLUTION EXPLICITE D'UN PROBLÈME DE TRANSPORT À COÛTS QUADRATIQUES ET APPLICATIONS

par

Michel Minoux*

CENTRE DE RECHERCHE SUR LES TRANSPORTS

CASE POSTALE 6128, SUCCURSALE A, MONTRÉAL, QUÉBEC, H3C 3J7

SOLUTION EXPLICITE D'UN PROBLÈME DE TRANSPORT À COÛTS QUADRATIQUES ET APPLICATIONS

par

Michel Minoux*

* Responsable du Département de Mathématiques appliquées
Centre national d'Etude des télécommunications
38-40, rue du Général Leclerc
92131 Issy-les-Moulineaux, FRANCE

Université de Montréal

Centre de recherche sur les transports - Publication #311
Département d'informatique et de
recherche opérationnelle - Publication #484

Juin 1983

S O M M A I R E

Le problème de transport à coûts quadratiques est un cas particulier du problème de flot à coûts quadratiques séparables : il est donc soluble en temps polynomial (MINOUX, M. 1982). Les applications en sont nombreuses par exemple dans le domaine de la prévision des matrices de trafic en Télécommunications, ou en économie, pour l'approximation des matrices input-output. On montre ici comment le problème peut être résolu de façon explicite lorsque les flots ne sont pas astreints à être non négatifs. On suggère ensuite l'utilisation de cette propriété pour résoudre efficacement le problème avec contraintes de non négativité par une méthode de dualité.

- - - - - - - - - - - - - - -

1) Position du problème

Etant donnée une matrice $\overline{X} = (\overline{x_{ij}})$ $i = 1,\ldots n$ $j = 1,\ldots n$

et 2n nombres l_i $(i = 1,\ldots n)$ et c_j $(j = 1,\ldots n)$ vérifiant

$$\sum_{i=1}^{n} l_i = \sum_{j=1}^{n} c_j = \sigma$$

on cherche à déterminer une matrice $X = (x_{ij})$ $i = 1,\ldots n$ $j = 1,\ldots n$

vérifiant :

$$\sum_{j=1}^{n} x_{ij} = l_i \quad (i = 1,\ldots n)$$

$$\sum_{i=1}^{n} x_{ij} = c_j \quad (j = 1,\ldots n)$$

$$x \geq 0$$

et approximant au mieux (au sens des moindres carrés) la matrice donnée $\overline{X}$. Ce problème se rencontre en économie (approximation des matrices input - output) et dans le domaine de la prévision des matrices de trafic dans les réseaux de communications (trafic routier, trafic téléphonique etc.). (cf. BACHARACH 1965, ALLEN et GOSSLING 1975, MINOUX 1978, DEBIESSE et MATIGNON 1980). Il s'agit d'un programme quadratique à contraintes linéaires qui appartient à la classe des problèmes suivants :

$$(P) \quad \begin{cases} \text{Minimiser} \quad \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{1}{2} (x_{ij} - \overline{x}_{ij})^2 \\ \text{sous les contraintes :} \\ \sum_{j=1}^{n} x_{ij} = l_i \quad (i = 1,\ldots n) & (1) \\ \sum_{i=1}^{n} x_{ij} = c_j \quad (j = 1,\ldots n) & (2) \\ x_{ij} \geq 0 \quad (\forall i = 1,\ldots n \ , \ j = 1,\ldots n) & (3) \end{cases}$$

Comme les contraintes (1), (2) et (3) sont celles d'un problème de transport, (P) peut être considéré comme un problème de transport à coûts quadratiques.

Pour le résoudre, un certain nombre de méthodes ont été proposées dans la littérature récente par exemple :

- la technique de relaxations successives par blocs (COTTLE & PANG 1981),
- la méthode de projections successives (BACHEM & KORTE 1980),
- la spécialisation de l'algorithme des pivots complémentaires (WOLFE 1959, LEMKE 1982, JAUMARD & MINOUX 1982).

Le problème (P) peut également être vu comme un cas particulier du problème de flot à coûts quadratiques séparables et peut donc être résolu en temps polynomial par une généralisation de l'algorithme out-of-kilter (cf MINOUX 1982).

Nous proposons ici une approche différente de ce problème basée sur la théorie de la dualité qui transforme la résolution de (P) en la résolution d'une suite de problèmes dans lesquels les contraintes (3) (non-negativité des variables) sont relaxées (problèmes relaxés).

Comme nous allons le voir l'intérêt d'une telle approche réside dans le fait que la solution optimale des problèmes relaxés peut être obtenue de façon explicite. La valeur de la fonction duale ainsi que le gradient de la fonction duale en un point peuvent donc être calculés de façon extrêmement efficace, et l'optimisation du problème dual peut être effectué par des algorithmes de gradient classiques.

2) Résolution explicite du problème sans contrainte de non négativité

Considérons, en toute généralité, un problème sans contrainte de non négativité du type :

$$(PR)\quad \begin{cases} \text{Minimiser} \displaystyle\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{1}{2}(x_{ij} - \bar{x}_{ij})^2 - \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\, x_{ij} \\ \text{sous les contraintes} \\ \displaystyle\sum_{j=1}^{n} x_{ij} = l_i \quad (i = 1,\ldots n) \qquad (1) \\ \displaystyle\sum_{j=1}^{n} x_{ij} = c_j \quad (j = 1,\ldots n) \qquad (2) \end{cases}$$

("problème relaxé")

Le terme linéaire $\sum_i \sum_j w_{ij}\, x_{ij}$ sera utilisé dans le §4 pour mettre en oeuvre la méthode de relaxation lagrangienne.

Notons ici qu'une condition nécessaire et suffisante pour que le problème relaxé (PR) ait une solution est que $\sum_{i=1}^{n} l_i = \sum_{j=1}^{n} c_j$.

Dans la suite les notations suivantes seront utilisées :

$$\bar{l}_i = \sum_{j=1}^{n} x_{ij}$$

$$\bar{c}_j = \sum_{i=1}^{n} x_{ij}$$

$$\bar{\sigma} = \sum_{i=1}^{n} \bar{l}_i = \sum_{j=1}^{n} \bar{c}_j = \sum_{i=1}^{n}\sum_{j=1}^{n} \bar{x}_{ij}$$

$$\sigma = \sum_{i=1}^{n} l_i = \sum_{j=1}^{n} c_j$$

Comme dans (PR) il s'agit de minimiser une fonction convexe sous contraintes linéaires, il s'agit d'un programme mathématique convexe. Les conditions de Kühn et Tucker sont donc nécessaires et suffisantes.

D'autre part, la fonction à optimiser étant strictement convexe, le problème a une solution optimale unique.

Ecrivons les conditions de Kühn et Tucker en notant λ_i les multiplicateurs de Kühn et Tucker associés aux contraintes $\sum_{j=1}^{n} x_{ij} = l_i \quad (i = 1,\ldots n)$

et μj les multiplicateurs de Kühn et Tucker associés aux contraintes

$$\sum_{j=1}^{n} x_{ij} = c_j \quad (j = 1,\ldots n)$$

Les conditions de Kühn et Tucker s'écrivent alors :
une condition nécessaire et suffisante pour que $x = (x_{ij})$ soit solution optimale est qu'il existe λ_i, μ_j (non contraints en signe) tels que :

$$\begin{cases} (x_{ij} - \bar{x}_{ij}) - w_{ij} + \lambda_i + \mu_j = 0 \ (\forall\, i, j) & (4) \\ \sum_j x_{ij} = l_i \quad (i = 1,\ldots n) & (1) \\ \sum_i x_{ij} = c_j \quad (j = 1,\ldots n) & (2) \end{cases}$$

En sommant sur i les équations (3) on obtient :

$$\forall j : c_j - \bar{c}_j - \sum_i w_{ij} + \sum_i \lambda_i + n\,\mu_j = 0 \qquad (5)$$

En sommant sur j on obtient :

$$\forall i : l_i - \bar{l}_i - \sum_j w_{ij} + n\,\lambda_i + \sum_j \mu_j = 0 \qquad (6)$$

Supposons connue la valeur $z = \sum_j \mu_j$ à l'optimum. Les équations (6) permettent alors d'obtenir une expression de chaque λ_i en fonction de z.

$$\lambda_i = -\frac{1}{n}\,(l_i - \bar{l}_i - \sum_j w_{ij} + z) \qquad (7)$$

On en déduit donc en posant $\sigma = \sum_i l_i = \sum_j c_j$

et $\bar{\sigma} = \sum_i \bar{l}_i = \sum_j \bar{c}_j$ et $W = \sum_{ij} w_{ij}$

$$\sum_i \lambda_i = -\frac{1}{n} (\sigma - \bar{\sigma} - W) - z$$

On déduit alors des équations (5) les valeurs μ_j par :

$$\forall\, j : \mu_j = -\frac{1}{n} \left(c_j - \bar{c}_j - \sum_i w_{ij} - \frac{\sigma - \bar{\sigma} - W}{n} - z\right) \qquad (8)$$

Maintenant, on remarque que $\forall\, i,j$ la somme $\lambda_i + \mu_j$ ne dépend pas de la valeur de z (qui peut être choisie arbitrairement). Les valeurs optimales x_{ij} sont donc déterminées par (3) en remplaçant les valeurs λ_i et μ_j par leurs expressions (7) et (8) ce qui donne $\forall\, i,j$:

$$x_{ij} = \bar{x}_{ij} + w_{ij} + \frac{1}{n} (l_i - \bar{l}_i - \sum_j w_{ij} + z) + \frac{1}{n} (c_j - \bar{c}_j - \sum_i w_{ij} - \frac{\sigma - \bar{\sigma} - W}{n} - z$$

$$x_{ij} = \bar{x}_{ij} + w_{ij} + \frac{1}{n} (l_i - \bar{l}_i - \sum_j w_{ij}) + \frac{1}{n} (c_j - \bar{c}_j - \sum_i w_{ij}) - \frac{1}{n^2} (\sigma - \bar{\sigma} - W)$$

Les formules ci-dessus permettent donc de déterminer de façon extrèmement efficace, la solution optimale d'un problème relaxé.

3) Un exemple numérique

Illustrons ce qui précède sur le petit exemple numérique suivant :

$n = 4$

$$X = \begin{pmatrix} 5 & 6 & 9 & 4 & \overline{l}_i \\ 2 & 1 & 0 & 3 & 32 \\ 0 & 5 & 10 & 9 & 10 \\ 4 & 8 & 7 & 6 & 36 \\ & & & & 29 \\ 26 & 27 & 21 & 33 & \leftarrow \overline{c}_j \end{pmatrix}$$

$w_{ij} = 0 \ (\forall\ i,j)$

Les valeurs $\overline{l}_i$ et $\overline{c}_j$ sont :

$\overline{l}_1 = 24$; $\overline{l}_2 = 6$; $\overline{l}_3 = 24$; $\overline{l}_4 = 25$

$\overline{c}_1 = 11$; $\overline{c}_2 = 20$; $\overline{c}_3 = 26$: $\overline{c}_4 = 22$

En fixant arbitrairement $z = 0$ il vient :

$$\lambda_1 = -\frac{1}{4}(32 - 24) = -2$$

$$\lambda_2 = -\frac{1}{4}(10 - 6) = -1$$

$$\lambda_3 = -\frac{1}{4}(36 - 24) = -3$$

$$\lambda_4 = -\frac{1}{4}(29 - 25) = -1$$

Puis en remarquant que $\frac{\sigma - \overline{\sigma} - W}{n} = \frac{28}{4} = 7$

$$\mu_1 = -\frac{1}{4}(26 - 11 - 7) = -2$$

$$\mu_2 = -\frac{1}{4}(27 - 20 - 7) = 0$$

$$\mu_3 = -\frac{1}{4}(21 - 26 - 7) = +3$$

$$\mu_4 = -\frac{1}{4}(33 - 22 - 7) = -1$$

De la relation $x_{ij} = \overline{x}_{ij} - \lambda_i - \mu_j$ ($\forall$ i,j) on en déduit la solution optimale du problème :

$\mu_j \longrightarrow$ $\quad -2 \quad 0 \quad 3 \quad -1$

$$\left(\begin{array}{cccc|c} 5 & 6 & 9 & 4 & -2 \\ 2 & 1 & 0 & 3 & -1 \\ 0 & 5 & 10 & 9 & -3 \\ 4 & 8 & 7 & 6 & -1 \end{array}\right) \qquad X^* = \begin{pmatrix} 9 & 8 & 8 & 7 \\ 5 & 2 & -2 & 5 \\ 5 & 8 & 10 & 13 \\ 7 & 9 & 5 & 8 \end{pmatrix}$$

$\uparrow$ λ_i

4) Résolution du problème avec contraintes de non négativité par une méthode duale

Revenons maintenant au problème initialement posé :

$$(P)\quad \begin{cases} \text{Minimiser } \sum_{ij} \frac{1}{2} (x_{ij} - \overline{x_{ij}})^2 \\ \text{Sous les contraintes} \\ \sum_j x_{ij} = l_i \quad (i = 1,\ldots n) & (1) \\ \sum_i x_{ij} = c_j \quad (j = 1,\ldots n) & (2) \\ -x_{ij} \leq 0 \quad (\forall\, i,j) & (3) \end{cases}$$

On va utiliser le fait que l'on sait résoudre de façon efficace (on a montré au § 2 comment obtenir une solution explicite) le problème (PR) sans contrainte de positivité sur les variables pour résoudre le problème (P) par une méthode de relaxation lagrangienne.

On associe aux contraintes (3) des variables duales $w_{ij} \geq 0$ $(\forall\, i,j)$.

En tout point $w = (w_{ij})$ $i = 1,\ldots n$, $j = 1,\ldots n$

La fonction duale $\phi(w)$ sera obtenue en résolvant le problème :

$$\phi(w) = \text{Minimiser } \sum_{ij} \frac{1}{2} (x_{ij} - \overline{x_{ij}})^2 - \sum_{ij} w_{ij}\, x_{ij}$$

Sous les contraintes

$$\sum_j x_{ij} = l_i \quad (i = 1,\ldots n)$$

$$\sum_i x_{ij} = c_j \quad (i = 1,\ldots n)$$

qui est un problème (sans contraintes de positivité) du type (PR).

Si $(x_{ij}^{*}(w))$ $i = 1,\ldots n$, $j = 1,\ldots n$ dénote la solution optimale (unique) de ce problème on sait que :

- la fonction duale $\phi(w)$ est concave, différentiable
- le vecteur de composantes $(x_{ij}^{*}(w))$ est le gradient de ϕ au point w.

On pourra donc résoudre le problème dual de (P') :

$$(D) \quad \begin{cases} \text{Maximiser } \phi(w) \\ \text{Sous les contraintes} \\ w \geq 0 \end{cases}$$

par une méthode de gradient (plus forte pente, gradient conjugué ou quasi Newton).

A l'optimum w^* du problème dual, le couple $(x^*(w^*), w^*)$ forme un <u>point col</u> et $x^*(w^*)$ est la solution optimale (unique) du problème (P).
Pour une justification de ces résultats et pour un exposé approfondi sur la théorie de la dualité en Programmation Mathématique, on pourra se reporter à MINOUX (1983) chapitre 6.

<u>Remerciements</u> :

Le contenu de ce papier est tiré d'un exercice donné aux étudiants de l'ENSTA à Paris, et je remercie M. FLORIAN pour m'avoir encouragé à lui donner une diffusion plus large.

BIBLIOGRAPHIE

ALLEN R.I.G., GOSSLING W.F. (1975) : *"Estimating and projecting input - output coefficients"*
Input - Output publishing Company, London.

BACHARACH M. (1975) : Estimating nonnegative matrices from marginal data"
International Economic Review 6, pp 294 - 310.

BACHEM A., KORTE B. (1980):"Minimum norm problems over transportation polytopes"
Linear Algebra and its Appl., 31, pp 103 - 118.

COTTLE R.W., PANG J.S. (1980):"On the convergence of a block successive overrelaxation method for a class of linear complementarity problems"
Technical Report SOL 80-17, Dept. Operations Research, Stanford University.

DEBIESSE J.L., MATIGNON G. (1980) : "Comparison of different methods for the calculation of traffic matrices"
Annales des Télécommunications 35, pp 91 - 102.

FERLAND J.A. (1974) : "Minimum cost multicommodity circulation problems with convex arc costs"
Transportation Science 8, 4 pp 355 - 360.

FORD L.R., FULKERSON D.R. (1962) :*"Flows in Networks"*
Princeton University Press.

JAUMARD B., MINOUX M. (1982) : "Un algorithme pour les problèmes de flots à coûts minimum et fonction de coût quadratique"
Note interne du CNET.

KENNINGTON J.L., HELGASON R.V. (1980) *"Algorithms for Network Programming"*
Wiley.

LEMKE C.E. (1962) "A method of solution for quadratic programs"
Management Science 8, pp 442 - 455.

MINOUX M. (1978) "Un algorithme de sous-gradient pour l'extrapolation des matrices de trafic pour une norme quadratique"
(note interne du CNET, non publiée, citée dans DEBIESSE et MATIGNON 1980).

MINOUX M. (1982) : "A polynomial algorithm for minimum quadratic cost flow problems"
Présenté au Congrès "Netflow 1983" Pise (Italie) Mars 1983.

WOLFE P. (1959) : "The Simplex method for quadratic programming"
Econometrica, vol 27, n° 3, pp 382 - 398.